



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2017

The ATAD2 bromodomain binds different acetylation marks on the histone H4 in similar fuzzy complexes

Langini, Cassiano ; Caffisch, Amedeo ; Vitalis, Andreas

Abstract: Bromodomains are protein modules adopting conserved helix bundle folds. Some bromodomain-containing proteins, such as ATPase family AAA domain-containing protein 2 (ATAD2), isoform A, have attracted much interest because they are overexpressed in many types of cancer. Bromodomains bind to acetylated lysine residues on histone tails and thereby facilitate the reading of the histone code. Epigenetic regulators in general have been implicated as indicators, mediators, or causes of a large number of diseases and disorders. To interfere with or modulate these processes, it is therefore of fundamental interest to understand the molecular mechanisms by which epigenetic regulation occurs. Here, we present results from molecular dynamics simulations of a doubly acetylated histone H4 peptide bound to the bromodomain of ATAD2 (hereafter referred to as ATAD2A). These simulations revealed how the flexibility of ATAD2A's major loop, the so-called ZA loop, creates an adaptable interface that preserves the disorder of both peptide and loop in the bound state. We further demonstrate that the binding involves an almost identical average pattern of interactions irrespective of which acetyl mark is inserted into the pocket. In conjunction with a likely mechanism of electrostatically driven recruitment, our simulation results highlight how the bromodomain is built toward promiscuous binding with low specificity. In conclusion, the simulations indicate that disorder and electrostatic steering function jointly to recruit ATAD2A to the histone core and that these fuzzy interactions may promote cooperativity between nearby epigenetic marks.

DOI: <https://doi.org/10.1074/jbc.M117.786350>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-141572>

Journal Article

Published Version

Originally published at:

Langini, Cassiano; Caffisch, Amedeo; Vitalis, Andreas (2017). The ATAD2 bromodomain binds different acetylation marks on the histone H4 in similar fuzzy complexes. *Journal of Biological Chemistry*, 292(40):16734-16745.

DOI: <https://doi.org/10.1074/jbc.M117.786350>

The ATAD2 bromodomain binds different acetylation marks on the histone H4 in similar fuzzy complexes

Received for publication, March 15, 2017, and in revised form, July 21, 2017 Published, Papers in Press, August 10, 2017, DOI 10.1074/jbc.M117.786350

Cassiano Langini,¹ Amedeo Caflisch,¹ and Andreas Vitalis²

From the Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland

Edited by Eric R. Fearon

Bromodomains are protein modules adopting conserved helix bundle folds. Some bromodomain-containing proteins, such as ATPase family AAA domain-containing protein 2 (ATAD2), isoform A, have attracted much interest because they are overexpressed in many types of cancer. Bromodomains bind to acetylated lysine residues on histone tails and thereby facilitate the reading of the histone code. Epigenetic regulators in general have been implicated as indicators, mediators, or causes of a large number of diseases and disorders. To interfere with or modulate these processes, it is therefore of fundamental interest to understand the molecular mechanisms by which epigenetic regulation occurs. Here, we present results from molecular dynamics simulations of a doubly acetylated histone H4 peptide bound to the bromodomain of ATAD2 (hereafter referred to as ATAD2A). These simulations revealed how the flexibility of ATAD2A's major loop, the so-called ZA loop, creates an adaptable interface that preserves the disorder of both peptide and loop in the bound state. We further demonstrate that the binding involves an almost identical average pattern of interactions irrespective of which acetyl mark is inserted into the pocket. In conjunction with a likely mechanism of electrostatically driven recruitment, our simulation results highlight how the bromodomain is built toward promiscuous binding with low specificity. In conclusion, the simulations indicate that disorder and electrostatic steering function jointly to recruit ATAD2A to the histone core and that these fuzzy interactions may promote cooperativity between nearby epigenetic marks.

Bromodomains are protein domains adopting a conserved left-handed four-helix bundle fold (1). The four helices are usually called α Z (closest to the N terminus), α A, α B, and α C (closest to the C terminus; see Fig. 1*a*). These domains are able to bind post-translationally modified lysine residues (2, 3) (*i.e.* acetyl-lysine located on N-terminal histone tails) to read part of the so-called histone code (4–6). From a now large number of crystal structures (7, 8), the acetyl-lysine binding site emerges as a well-defined pocket with mainly hydrophobic character

and lined by a few highly conserved residues (9). Across this protein family, the surrounding regions vary, however, and the highly flexible ZA loop (1) plays a prominent role in this regard. This loop, from ~23 to 55 residues in length, has the potential to make the binding site, which also involves the BC loop, partially disordered. This ability was highlighted by molecular dynamics (MD)³ simulations (10).

As bromodomains have no other known functions, the reading of the acetyl-lysine marks is inconsequential beyond blocking access of competing binders. Indeed, bromodomains usually perform their role as part of large multidomain proteins. It is assumed that they help in the recruitment and assembly of multiprotein complexes that regulate transcription and/or modify (write, erase) the histone code itself (7, 11–13). Based on this role, it is useful to formulate some likely properties of their binding to acetyl-lysine marks on histone tails. 1) The binding should be such that the residence times of the fully assembled complex do not add lag to the regulatory pathways involved. The speed of regulation at the gene level implies residence times of seconds or less (14–18). Moreover, most acetylation marks in the tails are turned over in minutes to hours themselves (19, 20). This means that the binding of the histone tail to the bromodomain is likely to be weak with fast on- and off-rates. 2) Because histone tails are positively charged under physiological conditions, electrostatic steering (21–24) can be a viable mechanism to accelerate binding, in particular to aid in the formation of productive encounter complexes. 3) Histone tails are completely disordered in solution, which has been argued to enable fast binding rates (25, 26). Because the tails are not known to fold upon binding, it is a reasonable conjecture that their *in vivo* interaction partners are not completely rigid either (27, 28). 4) Transcriptional regulation is an inherently noisy system due to the low concentrations of many of the molecular players involved (29). Redundancy in function, which translates to low overall specificity (30), and ultrasensitivity to regulation (28) are both properties that can be conferred by the formation of disordered complexes between two solvent-accessible entities (*e.g.* loops and tails). The epigenetic code is one example of a regulatory system likely to use both properties to generate predictable outcomes. Consequently, it

This work was supported by Grant 149897 from the Swiss National Science Foundation (to A. C.). The authors declare that they have no conflicts of interest with the contents of this article.

This article contains supplemental Movie S1 and Figs. S1–S12.

¹ To whom correspondence may be addressed: Dept. of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. Tel.: 41-44-635-5568; Fax: 41-44-635-6862; E-mail: caflisch@bioc.uzh.ch.

² To whom correspondence may be addressed: Dept. of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. Tel.: 41-44-635-5568; Fax: 41-44-635-6862; E-mail: a.vitalis@bioc.uzh.ch.

³ The abbreviations used are: MD, molecular dynamics; Kac, acetylated lysine; H3K14ac, histone H3 Lys¹⁴ acetylation, H4K5ac and H4K12ac, histone H4 Lys⁵ and Lys¹² acetylation, respectively; H4(1–16)K5acK12ac, N-terminal segment 1–16 of histone H4 with acetyl marks on residues Lys⁵ and Lys¹²; H4(9–16)K12ac, segment 9–16 of histone H4 with acetyl mark on residue Lys¹²; RMSD, root mean square deviation; PDB, Protein Data Bank.

Table 1**RMSD values between crystal structures of ATAD2A**

All-atom values used the heavy atoms in the 123 residues that had no missing atoms in all respective crystal structures, and the C_α values are based on 130 C_α positions.

RMSD to 4QUU	3DAI	4TT2
All-atom (Å)	0.803	1.362
C_α only (Å)	0.417	0.838

has been difficult to rationalize the often wide impact of pharmaceutical interference with bromodomain-histone interactions (31).

Here we ask what the molecular details allowing for such a set of properties could be. To further our mechanistic understanding of the reading of histone acetylation marks, we resort to MD simulations of a model system (*viz.* the bromodomain of ATPase family AAA domain-containing protein 2 (ATAD2)). There is considerable interest in this bromodomain (ATAD2A) as it is overexpressed in many types of cancer and its expression level correlates with poor survival probability (32–35). Although many insights have been obtained from comparisons of crystal structures in the apo-form with those complexed with small inhibitors or N-terminal histone tail fragments (36, 37), these analyses remain largely “static.” Similarly, the search for potent inhibitors through fragment-based screening (38–41) has yielded impressive results, but no clear rationale for peptide binding affinity, specificity, and kinetics could be extracted.

Experimentally, the recognition of different acetylation marks by ATAD2A has been explored with different techniques, showing a preferred interaction with histone tails H3K14ac (33) and H4K5ac (34). A value of $K_d = 22 \mu\text{M}$ for the binding of H4K5ac was reported using isothermal titration calorimetry (36). Protein crystallographic (X-ray) studies have concentrated on H4K5ac and H4K12ac (37). As we show in [supplemental Fig. S1](#), there is a caveat with crystals of ATAD2A in that the space group can have a larger influence on protein conformation than the presence of a binder, which is partially explained by crystal contacts substituting the ligand. For example, the root mean square deviation (RMSD) of C_α atoms between structures 4TT2 (36) and 4QUU (37), which both have a fragment of the N terminus of H4 with an acetylation mark on Lys⁵ bound, is 0.8 Å, yet that between structures 4QUU and 3DAI (8), the latter of which is an apo-crystal, is only 0.4 Å (Table 1). 3DAI and 4QUU belong to the same space group ($P6_522$), which differs from that of 4TT2 ($P4_22_12$). Whereas crystal contacts probably modulate this effect, the small differences between the resolved X-ray structures of ATAD2A are often found in the apertures of the ZA loop, suggesting that conformational flexibility influences the binding of the peptide (36).

Previous MD simulations extended this suggestion and showed not only that the ZA loop is largely disordered (10) but also that alternative binding poses can be realized, for example, for the TAF1(2) bromodomain (42). In particular, a more buried conformation was observed, in which three of the conserved crystal waters found in most deposited structures are displaced. The latter study relied on truncated acetyl-lysine binders, and it constitutes another aim of our present study to investigate how likely alternative binding poses are in the context of a signifi-

cant portion of the N-terminal histone tail, which is likely to make specific interactions itself.

To elucidate details of the binding of histone tails to bromodomains further, we use MD simulations of ATAD2A bound to the N-terminal segment 1–16 of histone H4 with acetyl marks on residues Lys⁵ and Lys¹² (hereafter abbreviated as H4(1–16)K5acK12ac). These two lysine residues are robustly found to be acetylated on newly synthesized H4 (43), and the diacetylated and both monoacetylated peptides interact with ATAD2A (44). The residue numbering below will refer to the complete protein sequence (UniProt entry Q6PL18). Crystal structures 4QUU and 4QUT (37) offer two alternative starting conformations with either Kac5 or Kac12 inserted into the binding pocket (Fig. 1*b*). [Supplemental Movie S1](#) offers a detailed look at the Kac5-inserted structure and its surface properties. In the remainder of the paper, we compare the two binding modes and their impact on ATAD2A conformation with a particular focus on loop disorder. The peptide is stably bound during these holo-simulations, and we additionally analyze MD data of the unbound states of both ATAD2A (apo-simulations) and histone tail for comparison (see summary in Fig. 1*c* and under “Experimental procedures”). We conclude by discussing the implications of our results for the ability of bromodomains to read the histone code.

Results

Below, we first analyze the binding mode of the two histone marks (*viz.* H4(1–16)K5acK12ac with buried Kac5 or Kac12) and then investigate to what extent they influence the flexibility of the ZA loop with respect to each other and to the apo-form.

Bromodomain binding does not strongly restrict the peptide's conformational space

The distributions of the radii of gyration ([supplemental Fig. S2](#)) and two-dimensional histograms of radius of gyration *versus* asphericity (Fig. 2) show that the H4(1–16) peptide exhibits similar average behavior irrespective of whether it is bound and, when bound, irrespective of which Kac is inserted. The dominant states are always those with small values for both parameters, indicating that the peptide prefers to fold back onto itself. This is somewhat unexpected, given the sequence of the H4 tail, which is dominated by positively charged and glycine residues. The average radii of gyration are not different in a statistically reliable way ([supplemental Fig. S2](#)). Whereas the shape of the distributions for the unbound peptide indicates a slightly broader conformational envelope, this is a weak effect. It thus appears as if the peptide is solvated similarly well by either the protein or aqueous buffer. The most important implication of this is that the histone tail's disorder can be maintained in the bound state, which means that binding may be entropically neutral or even favorable if one also considers the contribution of the solvent.

The individual trajectories of the bound peptide are on the submicrosecond time scale (see Fig. 1*c*), and each one samples a restricted area of phase space as is apparent from [supplemental Fig. S2](#) and will become clear in general below. This is not the case for the simulations of the unbound peptide, where the S.E. values are much smaller. The differences in sampling quality

ATAD2A binds acetyl-lysine marks on H4 in fuzzy complexes

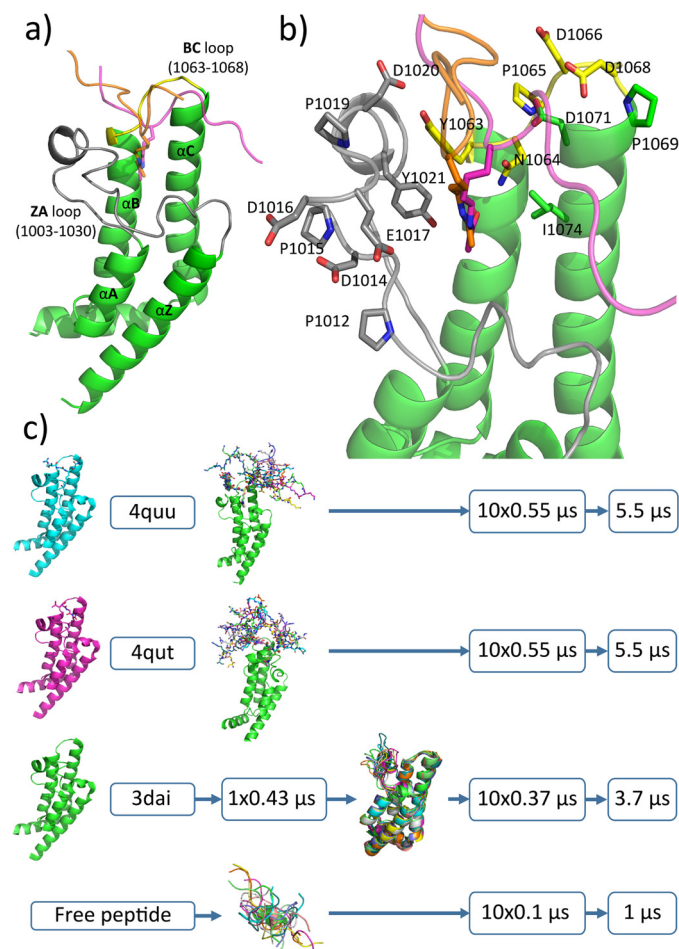


Figure 1. Schematic views of ATAD2A. *a*, the helix bundle architecture of the domain is shown in green with the four helices (A, B, C, and Z) labeled. The two important loops connecting α Z with α A (ZA loop; gray) and α B with α C (BC loop; yellow) are indicated. The (slightly truncated) backbone of the simulated peptides with Kac12 (orange) and Kac5 (pink) inserted is depicted along with a stick representation of this residue. In both cases, the peptide structure is taken from the parts resolved in 4QUU and 4QUT along with an example reconstruction for the remainder. The protein corresponds to the first stored snapshot from one of the Kac5-inserted simulations (see also supplemental Movie S1). *b*, we present a focused view of the interaction surface with a number of residues highlighted. Residues Tyr¹⁰⁶³ and Asn¹⁰⁶⁴ on the BC loop and Tyr¹⁰²¹ on the ZA loop form the immediate environment of the Kac side chain. The putative gatekeeper residue is Ile¹⁰⁷⁴. *c*, simulation inventory. Starting from crystals 4QUU and 4QUT, we added the residues missing in the ligand and generated 10 starting structures from each crystal by Monte Carlo sampling (see supplemental Methods). We started one simulation from the crystal 3DAI, from which we saved one snapshot every 20 ns starting at 120 ns, for a total of 10 structures. These were used as starting points of the additional apo-runs. Starting conformations for the unbound peptide runs were chosen such that they covered the envelope of the holo-runs in Fig. 2. The superposed schematic images of the peptides represent the actual starting conformations in all cases.

highlight that the rates of conformational fluctuations are significantly reduced for the bound peptide, although the envelope may be similar. We do emphasize that the similar average behavior extracted from the 10 trajectories in each case is unlikely to be mere coincidence.

In light of this, there is a conceptual caveat that we sample orientations of the truncated C terminus of the H4 peptide, which may well be incompatible with the presence of the full histone. To address this concern, we heuristically calculated the fraction of peptide-bound conformations in which the C-ter-

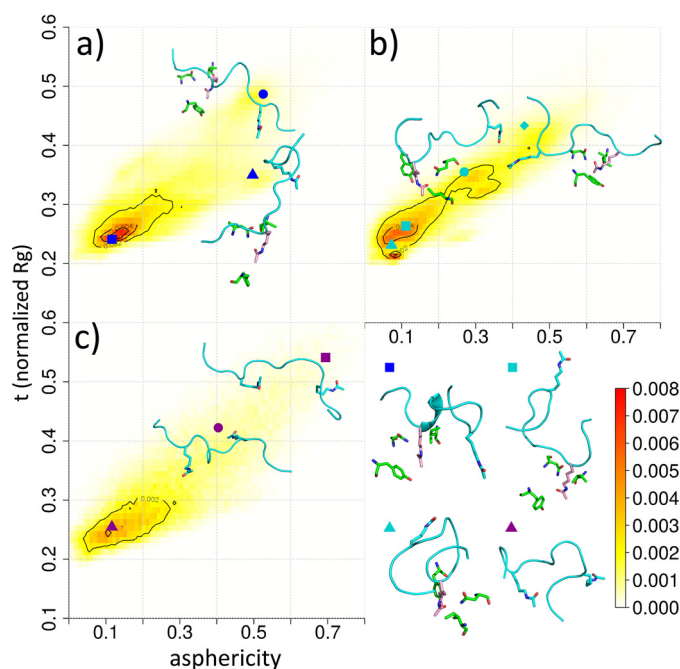


Figure 2. Two-dimensional histograms of parameter t (normalized radius of gyration) versus asphericity (measure of shape) for the peptide. Both measures are normalized to an interval of 0 to 1 and tend to correlate with each other. The color bar in the bottom right panel applies throughout. *a*, data for bound peptide with Kac5 inserted are shown. Images of specific snapshots are overlaid with the corresponding shape and size values indicated by different symbols and colors. In the schematic images, selected residues of the bromodomain (Tyr¹⁰²¹, Asn¹⁰⁶⁴, and Ile¹⁰⁷⁴) are shown as green sticks, and the Kac tail is shown in either pink (inserted) or cyan (not inserted) sticks. *b*, same as *a* for bound peptide with Kac12 inserted. *c*, same as *a* for the free peptide in aqueous solution. Here, the first half of each continuous trajectory was discarded because the initial states of the peptide (Fig. 1c) may not be representative for the unbound state. Four schematics referring to *a*, *b*, *b*, and *c* are displaced to the bottom right.

minal *N*-methyl amide group is buried (see supplemental Fig. S3). The fractions of nonrepresentative conformations calculated this way are 8% (Kac5 inserted) and 13% (Kac12 inserted), respectively. Although obviously an approximation, this suggests to us that the observed peptide conformations are fair representatives also in the context of the entire histone. In addition, we tested the hypothesis that a change in secondary structure content of the histone tail is masked by a lack of shape and size changes. However, regular elements of secondary structure are almost never observed in the simulations (supplemental Fig. S4), which is expected because 8 of 16 peptide residues are glycines. The invariance in Fig. 2 and supplemental Fig. S2 is not an inherent property of the sequence, however, and it has been shown that the addition of acetylation marks does have a significant influence on the free energy landscape (45).

The binding is characterized by a consensus mean interaction pattern for both acetyl marks

The peptide *versus* ATAD2A and peptide *versus* peptide residue-level contact maps in Fig. 3 highlight that the mean interaction profile is largely identical if the inserted Kac residue is used as a reference point along the sequence of the histone tail. The S.E. values for these data per cell are shown in supplemental Fig. S5. Supplemental Fig. S6 reveals that the interaction profiles are indeed nearly superimposable if peptide residues

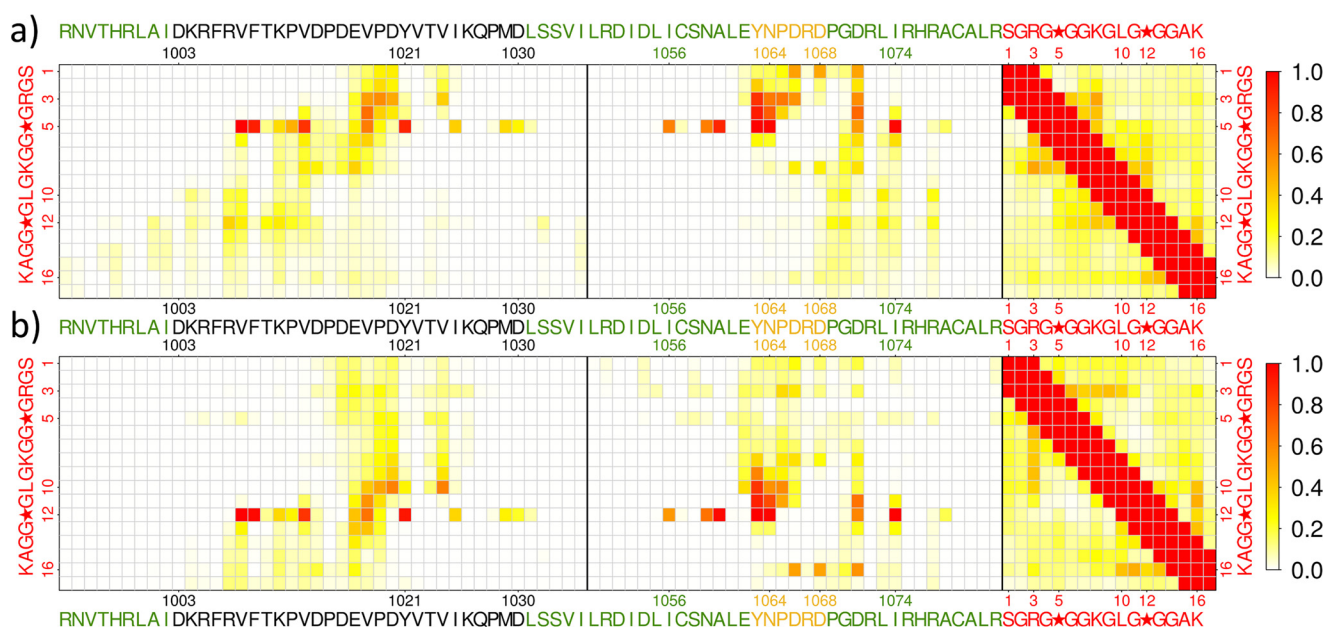


Figure 3. Contact maps for the two sets of holo-simulations. Occurrence frequencies for interresidue contacts are plotted for contacts between the peptide (vertical axis) and relevant parts of the bromodomain as well as the peptide itself (horizontal axis). A contact was counted whenever any pair of atoms was within 5.0 Å. The color of the one-letter residue labels identifies peptide (red), ZA loop (black), BC loop (orange), and helix (green) residues. The star symbol denotes Kac. *a*, data for the simulations with Kac5 inserted; *b*, data for the simulations with Kac12 inserted.

are shifted accordingly. This is not completely unexpected, given that the sequence context is similar (*i.e.* $^1\text{SGRGKacGKG}^9$ and $^8\text{KGLGKacGGAK}^{16}$). The largest differences are encountered where residue properties differ most (*i.e.* Arg³ versus Leu¹⁰). It is important to point out that many of the contextual interactions are not present simultaneously (*i.e.* the pattern results from the disordered ensembles of peptide and ZA loop, which retain their disorder upon binding (see below)). This structural ambiguity is a hallmark of fuzzy complexes (46), and it here involves two disordered binding partners.

The intermolecular contact patterns in Fig. 3 allow the following observations. First, the buried Kac is always involved in the largest number of contacts with ATAD2A, and these include the expected specific interactions with Asn¹⁰⁶⁴, Tyr¹⁰²¹, and the gatekeeper, Ile¹⁰⁷⁴ (Fig. 1*b*). Second, the contacts between the residues immediately adjacent to the inserted Kac and the residues in the ZA and BC loops (residues Asp¹⁰⁰³–Asp¹⁰³⁰ and Tyr¹⁰⁶³–Asp¹⁰⁶⁸, respectively) are preserved (*i.e.* no reversal of peptide orientation occurs). Third, the following up/downstream residues with a positive charge (Ser¹/Lys⁸ and Lys⁸/Lys¹⁶ for Kac5 and Kac12, respectively) are involved in favorable electrostatic interactions with the aspartate triad of the BC loop (*viz.* Asp¹⁰⁶⁶, Asp¹⁰⁶⁸, and Asp¹⁰⁷¹). There appears to be an inversion of contact preference for this triad, and it seems that Kac5 insertion favors the upstream position (amino group of Ser¹ over Lys⁸), whereas Kac12 insertion favors the downstream position (Lys¹⁶ over Lys⁸). A possible reason is that both of the favored residues are terminal in the construct we used (excluding the C-terminal cap), but this result may not be significant (see supplemental Fig. S5).

Even the specific binding motif of the acetyl mark is subject to distinct subpopulations, as shown in supplemental Fig. S7, which resolves two major states differing in whether Asn¹⁰⁶⁴ or Tyr¹⁰²¹ form a direct hydrogen bond to the oxygen atom of the

acetyl mark. The typical conformation found in crystals (direct hydrogen bond to Asn¹⁰⁶⁴) (9) is well-populated in the simulations irrespective of which Kac is inserted. Importantly, neither Fig. 3 nor supplemental Fig. S6 reveal dramatic differences between the two holo data sets, and the relative deviations in the superimposed contact patterns (supplemental Fig. S6) never exceed ~50%.

The main results are thus that the interaction surface covers a significant portion of the bromodomain, that the average contact patterns are preserved irrespective of the inserted Kac side chain, and that preferential interactions exist to histone peptide residues as far as four positions away from the acetyl mark. Because many of the contacts do not occur simultaneously, it is not possible to conclude that ATAD2A recognizes a particular sequence motif. Instead, it seems as if the complex must be viewed as a broad ensemble of bound states. In this view, sequence context is likely to contribute by coarser parameters, such as glycine content or the rough spacing of positive charges, to both the preferred orientation of binding and to net affinity, which remains moderate (36).

The binding interface is partially disordered in both bound and unbound states

Fig. 3 shows that the histone peptide makes significant interactions with ATAD2A's ZA loop (Fig. 1*b*, residues 1003–1030) as well as the BC loop and its surroundings (residues 1056–1074). Across space groups, crystallographic B factors provide conflicting information regarding the residual heterogeneity in the crystalline state (supplemental Fig. S1*b*). We hypothesize that the ZA loop is the major site of disorder, and the coordinate-based root mean square fluctuation profiles of the C_α atoms in supplemental Fig. S8 seem to support this claim. However, the root mean square fluctuation values also implicate the BC loop region and both termini as flexible. It is a caveat that

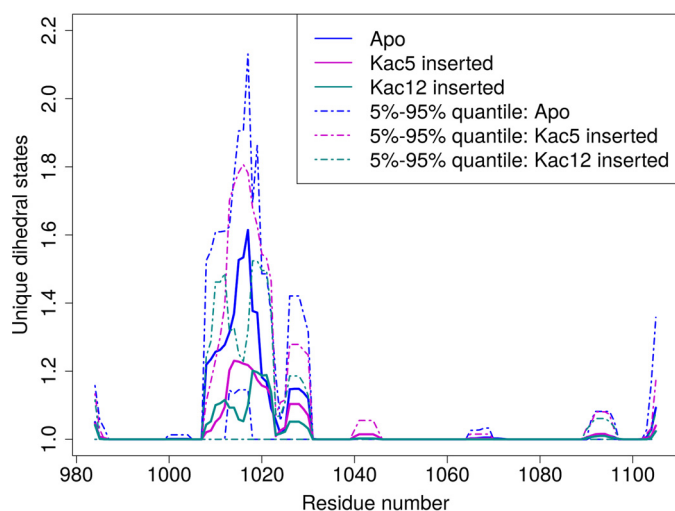


Figure 4. Conformational heterogeneity across the protein sequence. Backbone dihedral angles were converted to a simple 3-state model (see supplemental Fig. S9). Blocks of 5 residues were used to construct vectors of length 5 for the central residue, giving a maximum of $3^5 = 243$ states. The average number of unique vectors across overlapping time windows of 5 ns was calculated, and this average is plotted along with the 5 and 95% quantiles for block averages across 30 (both holo cases) and 22 blocks (apo), respectively. Note that side-chain dihedral angle states are ignored in this analysis. The number of discovered states for proline residues is probably overestimated due to the presence of a favored intermediate state between the α and polyproline regions, which we did not account for (see supplemental Fig. S10).

these values are influenced by the requirement to align structures, which is difficult to do in a way that treats all segments of the protein equally. We thus show in Fig. 4 a complementary measure of heterogeneity based on unique strings of backbone dihedral angles across 5-residue segments and 5-ns intervals. This analysis assumes a 3-state model per residue (see legend to Fig. 4 for details). Fig. 4 gives a result that is qualitatively consistent with a visual inspection of the trajectories, which, aside from the ZA loop, reveals ATAD2A to be surprisingly rigid. The most disordered stretch is the same in both the holo- and apo-forms (residues ~ 1008 – 1022) and coincides exactly with the dominant interaction surface on the protein involving ZA loop residues (Fig. 3). This is a stunning result as it means that the ZA loop retains most of its disorder upon binding of the histone tail. This type of fuzzy complex (28, 46) is likely to facilitate fast binding as no disorder-to-order transition occurs. Furthermore, the segment from residue 1025 to 1030 loses some heterogeneity in the holo-forms, which is surprising, given that it does not often form direct contacts with the peptide (Fig. 3). Last, Fig. 4 reveals the rest of the bromodomain to be essentially free of significant (anharmonic) conformational fluctuations at the backbone level.

We thus conclude that partial disorder for both partners and binding are not mutually exclusive even when no disorder-to-order transition occurs. This is somewhat similar to prior observations of the disordered binding of small molecules to a disordered peptide (47). The 3-state model for backbone dihedrals (supplemental Fig. S9) also allows us to quantify the dynamics of conformational transitions in the bound partners. Because the Kac anchor has to be inserted, it is presumably beneficial for transient binding if the surrounding sequence context does not add slow steps to the (un)binding. Although

we do not observe spontaneous (un)binding events, Fig. 5 depicts the total transition rates for each residue in the peptide and in the loop regions in the bound and unbound states. On average, rates for the peptide are roughly 1 order of magnitude faster than for ATAD2A (nanosecond time scale). As implied by Fig. 4, the BC loop as well as the N-terminal section of the ZA loop do not undergo any backbone conformational transitions.

Fig. 5b highlights a few ZA loop residues displaying major deviations between apo- and holo-forms within the structure. For some of these (e.g. Thr¹⁰¹⁰, Pro¹⁰¹⁵, and Glu¹⁰¹⁷) values in the apo-runs are comparable with those of peptide residues, highlighting 1) the disordered nature of the ZA loop and 2) that the presence of the histone peptide is of course not entirely inconsequential. The segment Pro¹⁰¹⁵–Asp¹⁰²⁰ in particular forms the highest peak in Fig. 4. As will become clear below, it tends to extend much more into solution in the apo-form, and this constitutes a substantial departure from the crystal structure. Glu¹⁰¹⁷ engages in salt bridges with Lys⁸ of the peptide (Fig. 3), which is likely to cause the slowdown of its dynamics in the holo-runs. Last, Asp¹⁰²⁰ makes it clear that a disordered binding partner can even enhance the rate of spontaneous fluctuations of the protein for a residue in direct contact (Fig. 5b).

Holo-trajectories return to the crystallographic structures more often than apo-trajectories

To quantify the mutual similarity and level of overlap between the apo- and holo-runs, we turn to a powerful analysis and visualization strategy called a SAPPHERE plot (48). Briefly, all snapshots are arranged by geometric criteria that use as the only input a definition of pairwise distance between conformations. The resultant sequence (termed progress index) is annotated with structural information, times of occurrence (kinetic trace), and a cut function able to highlight basins and barriers (see “Experimental procedures” and Ref. 49 for details).

Fig. 6 shows a SAPPHERE plot constructed with the coordinate RMSD of the N, O, and C _{β} atoms of residues of ATAD2A’s ZA and BC loops (see supplemental Methods for the full list) while aligning on the C _{α} atoms of the four α -helix bundle. Focusing first on the cut function (solid lines), a number of large and many smaller metastable states are easily identifiable. The dot pattern gives the sampling times with individual trajectories offset from each other. It is clear that most basins are predominantly sampled by either apo- or holo-runs but not both. The most populated metastable states are visited multiple times (recurrently), which is in contrast to many of the smaller ones. Whereas there are a few states unique to individual holo-trajectories (e.g. around progress index values of 4×10^5), no major state emerges for which the following holds: It is sampled recurrently within the set of trajectories with Kac5 inserted but is not visited when Kac12 is inserted (and vice versa). This suggests that, aside from individual trajectory divergence, the holo-simulations are a relatively homogeneous set that differs clearly from the apo data set.

This result is corroborated by the joint clustering of all snapshots (see “Experimental procedures”). We obtained 4251 clusters of size 20 or more representing 93% of all snapshots. There are 748 among these to which both holo-conditions contribute at least 10% of the snapshots, and this includes 51 of the 100

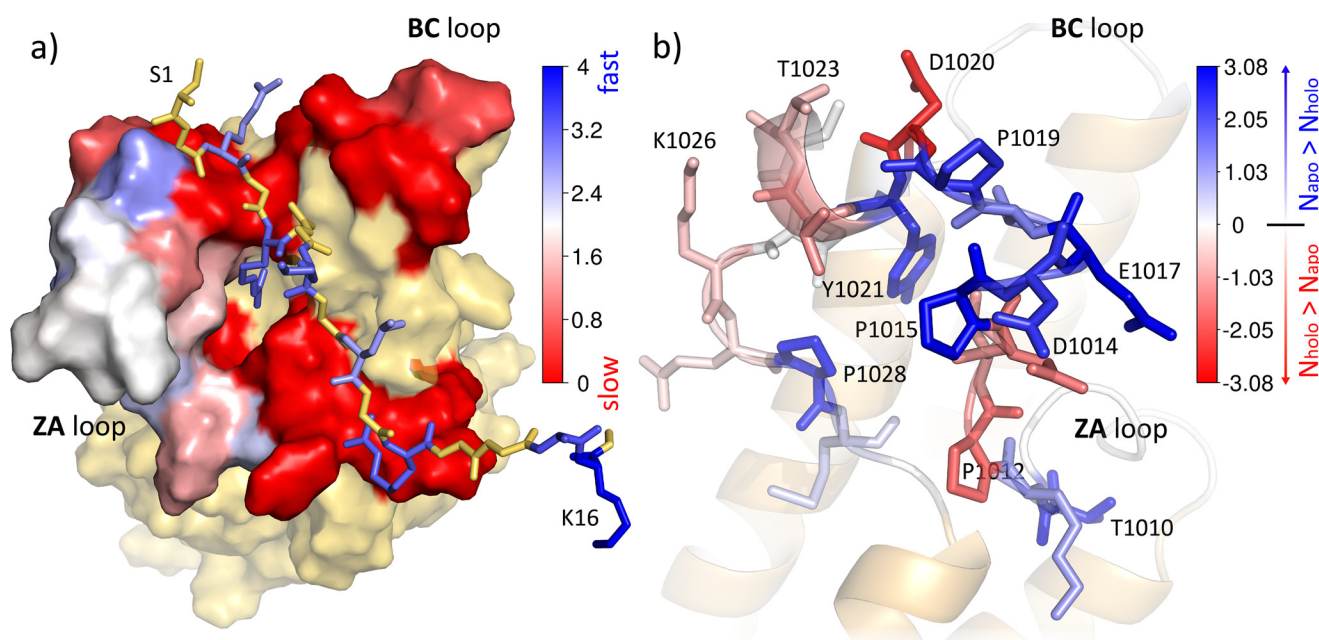


Figure 5. Visualization of transition rates between regions of the Ramachandran plot for backbone dihedrals of peptide and loop regions. Residues not taken into account for the calculation are in yellow, which includes all glycines. Selected residues and the two loops are highlighted by labels. *a*, transition rates for both the histone peptide and the ZA and BC loops in all holo-trajectories. Residues are shown in stick (histone peptide) and surface (ATAD2A) representations, respectively, and colored according to the total number of transitions in logarithmic scale (see supplemental Figs. S9 and S10). The (static) conformation is taken from the first snapshot of the first trajectory with Kac5 inserted. *b*, difference in transition rates between apo- and holo-states for prominent loop residues. Loop residues are shown as sticks along with the rest of the protein as schematics. Note that the viewing angle is different relative to *a*. Here, the color scheme is the logarithm of the absolute magnitude of the difference in numbers of transitions, and blue and red colors indicate a greater number of transitions in apo- and holo-runs, respectively.

largest clusters. Conversely, the overlap numbers for Kac5 inserted with apo and Lys¹² inserted with apo are 139 and 189, respectively, and here the shared clusters are never among the 100 largest overall. A high level of separation between apo- and holo-runs may be suggestive of a binding mechanism tending more toward the classical “induced fit” model, but this inference is weakened by the fact that divergence of individual trajectories does occur (see supplemental Figs. S2 and S5).

Up to progress index values of $\sim 3.2 \times 10^5$, Fig. 6 predominantly shows basins sampled recurrently in holo-runs irrespective of which Kac is inserted. Both DSSP (50) and dihedral angle annotations reveal these states to differ from each other in the joint state of ZA loop residues Pro¹⁰¹⁹ and Asp¹⁰²⁰ (see annotation). The snapshots from the apo-runs are also concentrated in a recurrent sampling region, which is located at 4.8×10^5 to 6.4×10^5 . These basins appear quite homogeneous, and there are clear differences between basins (e.g. again in residues Pro¹⁰¹⁹ and Asp¹⁰²⁰). Supplemental Fig. S11 uses histograms of data in annotations in Fig. 6 to illustrate that heterogeneity in these two residues occurs similarly in all three simulation sets. Importantly, many other features deviate strongly from all three crystal structures, and the higher similarity of holo- than apo-runs to crystal structures is best illustrated using the small segment Pro¹⁰¹⁵–Glu¹⁰¹⁷, where the tight turn is consistently lost in the apo-runs (DSSP assignment is empty or “bend”). Indeed, supplemental Fig. S11 clearly shows a split between apo- and holo-runs for Asp¹⁰¹⁶.

Different states can have different binding apertures

The analyses in Figs. 4 and 5 revealed the stretch of residues 1015–1020 in the ZA loop to be the most flexible one in the

entire domain. There is a cluster of closely spaced negative charges (Asp¹⁰¹⁴, Asp¹⁰¹⁶, Glu¹⁰¹⁷, Asp¹⁰²⁰), and Glu¹⁰¹⁷ and Asp¹⁰²⁰ engage in frequent contacts with a number of residues of the peptide in the holo case (Fig. 3). In the absence of the peptide, the structure opens up to become more solvent-exposed. This reshaping of the ZA loop is documented by an annotation with selected intramolecular distances in Fig. 6 (bottom of structural annotations). Distances between C α atoms of the ZA loop and the gatekeeper residue, Ile¹⁰⁷⁴, are shown. For example, the aforementioned localized dihedral change of Pro¹⁰¹⁹ and Asp¹⁰²⁰ in the holo case coincides with a small but noticeable difference in the distance of Glu¹⁰¹⁷–Ile¹⁰⁷⁴. For the most homogeneous apo-basin located at progress index values of $\sim 6 \times 10^5$, the same distance adopts large values not encountered in holo-simulations, which is revealed clearly by histograms (supplemental Fig. S11). This contrast between apo- and holo-states indicates that peptide binding tends to pull the ZA loop in to recruit additional binding interfaces, which is intuitive. Consistent with the disorder of the peptide–protein interaction, not all bound states show a fully contracted ZA loop (e.g. the basins near 3.3×10^5 and 6.6×10^5).

Fig. 7 shows a juxtaposition of two conformations that are either “closed” or “open” according to this simplified definition. Taken together, Figs. 6 and 7b provide evidence for the claim made previously that the segment Pro¹⁰¹⁵–Asp¹⁰²⁰ within the ZA loop can extend far into solution. From Fig. 6 and supplemental Fig. S11, it is also clear that this happens predominantly, but not exclusively, in the apo-state. The suggestion that this provides a mechanism of increasing the on-rate for histone tail

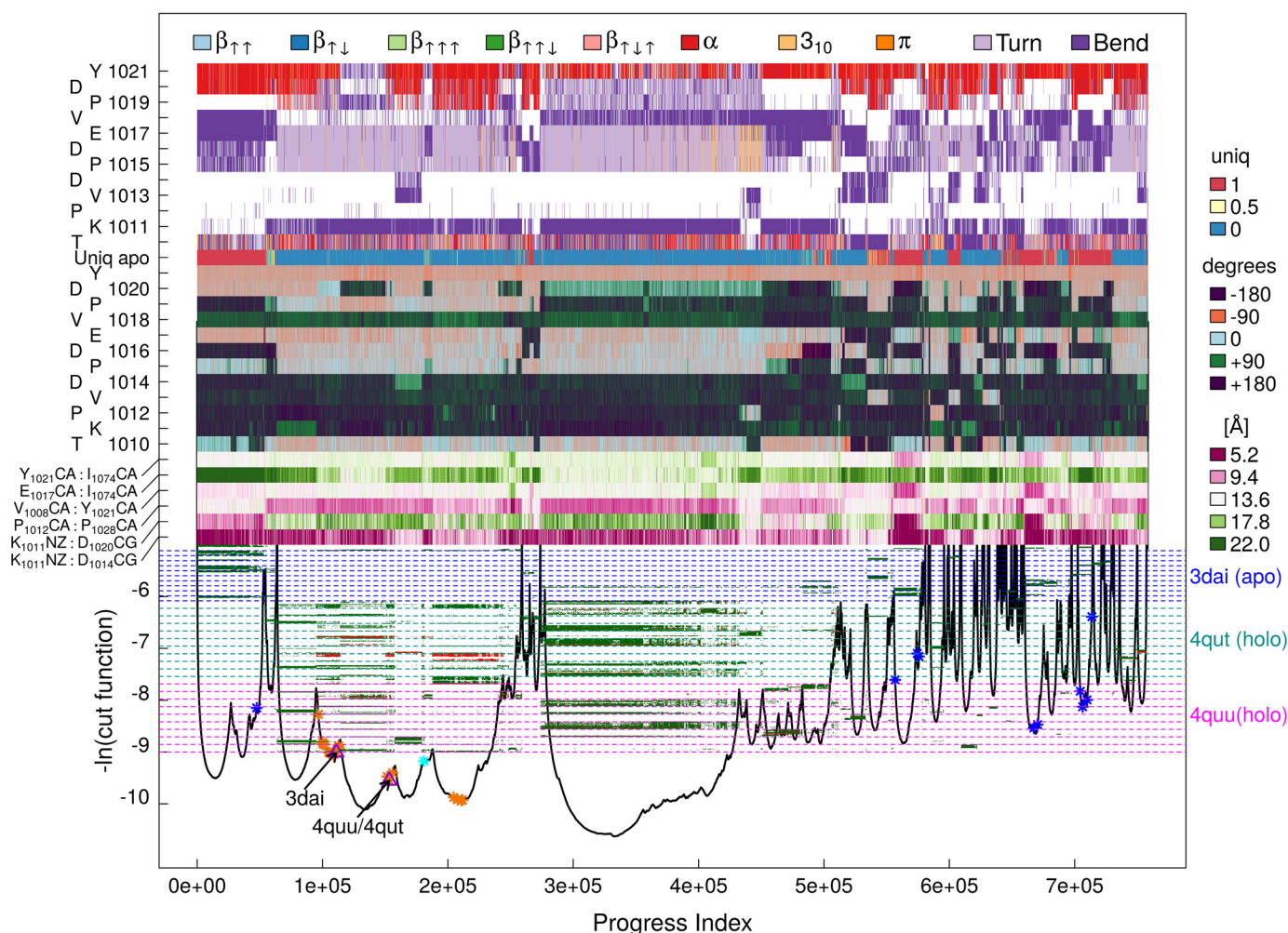


Figure 6. SAPPHERE plot of holo- and apo-trajectories. The distance metric used is the RMSD metric described under "Experimental procedures." The bottom shows a 3-state cut profile (49) (local partition width of 70,000 snapshots) highlighting the various free energy basins and barriers. Colored symbols locate the first recorded points of each trajectory (orange and blue/cyan symbols for holo and apo, respectively) and the underlying crystal structures (labeled). Note that a large number of holo starting points are superposed near the location of the crystal structures as expected. The dot pattern plots the time of occurrence for each independent run. Time increases from bottom to top, and multiple runs are offset on the y axis, as indicated by the colored dashed lines, which denote the different simulation sets. Green dots, conformations compatible with histone tail extension; red dots, other conformations (see supplemental Fig. S3). The color patterns stack, from bottom to top, various annotations per snapshot: distances, backbone ψ -angles (one-letter code), cluster uniqueness, and DSSP assignment (also one-letter code). The labels on the left explain what is plotted in each row. Color legends are found on top (DSSP) and to the right (remainder). Every snapshot is part of a cluster, $c(i)$, and cluster uniqueness is defined as $n_{c(i)}^{\text{apo}}/n_{c(i)}^{\text{total}}$, where n is the number of snapshots and $n_{c(i)}^{\text{apo}} + n_{c(i)}^{\text{holo}} = n_{c(i)}^{\text{total}}$.

binding is obvious, in particular if we consider the net negative charge of this segment.

The analysis is robust

One may be tempted to speculate that the split in Fig. 6 into so many distinct states for a domain that remains folded throughout is an artifact of the properties and dimensionality of the metric we use. Here, we provide evidence against this notion. First, the three crystal structures (purple) superimpose and co-localize with many of the first snapshots of the holo-runs (orange) and the first snapshot of the first apo-trajectory (cyan) but not with any first snapshots of the 10 apo-runs started from MD snapshots (dark blue). This matches the expected memory of the initial conditions we used. With an uninformative distance function, we would expect these points to be scattered across the plot. Second, we obtain a generally good partitioning of dihedral angle states in Fig. 6, although the metric is based on RMSD. For comparison, supplemental Fig.

S12 presents a complementary SAPPHERE plot from the data using the dihedral angles in the ZA and BC loops to define distance (see the supplemental material for the full list). This plot demonstrates that most of the snapshots from basins delineated in Fig. 6 also form basins in this new representation, thus confirming the robustness of the partitioning. Third, we show in Fig. 8 explicit projections of reduced dimensionality.

Fig. 8 (a–c) uses two geometric distances between the gatekeeper and ZA loop residues (*viz.* Glu¹⁰¹⁷–Ile¹⁰⁷⁴ and Tyr¹⁰²¹–Ile¹⁰⁷⁴). Despite the dimensionality now being 2 instead of 297 as in Fig. 6, the holo- and apo-runs sample only a marginal overlap region away from the respective minima of the projected energy surface (Fig. 8c). Conversely, the two holo-sets largely overlap. Consistent with the prior analysis, the main basin of the two holo-runs but not the apo-runs includes the crystal structures (Fig. 8, a and b). To further corroborate the results derived from Fig. 6, Fig. 8d utilizes a projection of the snapshots onto the first two principal components of the

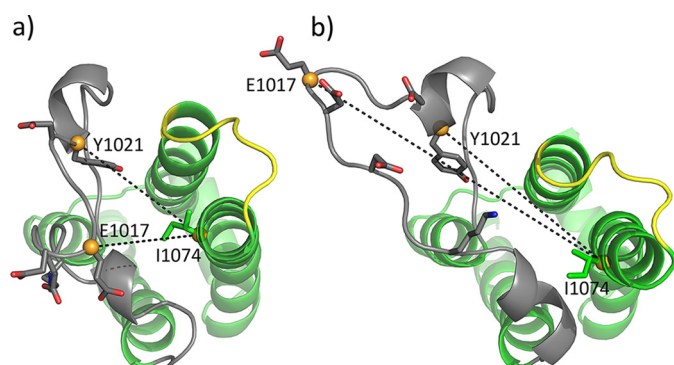


Figure 7. Illustration of changes to binding site aperture. *a*, a closed state from the main basin of the holo-simulations (peptide omitted). The domain is shown as *schematics* with the acidic residues mentioned in “Results” highlighted as sticks along with Lys¹⁰¹¹, Tyr¹⁰²¹, and Ile¹⁰⁷⁴. The coloring of the schematic is the same as in Fig. 1 (*a* and *b*). The aperture distances (black dashed lines) shown in Fig. 6 and supplemental Fig. S12 use the reference points indicated by the orange spheres. These residues are additionally identified with text labels. *b*, an open state from the main basin of the apo-simulations. Note how far Glu¹⁰¹⁷ can extend into solution. The C_α RMSD between the two conformations is 3.6 Å when including the ZA loop and 1.2 Å when ignoring it.

dihedral angles of the ZA loop. Again, the three sets of simulations appear to reside in overlapping regions only for the two holo data sets. We are therefore confident that there is a robust difference in the states sampled by apo- and holo-runs. This is in stark contrast to the negligible differences between the holo-runs with either Kac5 or Kac12 inserted.

Discussion

We simulated, using explicit solvent canonical MD, the complex of the bromodomain of ATAD2A with a diacetylated N-terminal tail of histone H4, specifically H4(1–16)K5acK12ac, in two different binding modes (*i.e.* with buried Kac5 and Kac12). Together with simulations of the unbound protein and peptide, the following picture emerged. On the time scale of the single simulation runs ($\sim 0.55 \mu\text{s}$), we did not observe any unbinding events or interconversions between the two acetylated marks. The two binding modes involve an almost identical pattern of interactions with the bromodomain (Fig. 3 and supplemental Fig. S6), which implies that they prefer the same orientation. The peptide, as expected based on sequence, does not fold upon binding (51); nor does it exhibit polymeric properties that are strongly altered (Fig. 2 and supplemental Fig. S2). Peptide binding has a mild influence on local heterogeneity (Fig. 4). As expected, rates of interconversion of backbone dihedral angles are slowed down for both binding partners in the bound state (supplemental Figs. S2, S5, and S10). There are no obvious and statistically meaningful differences in the sampled protein conformational states of the two different binding modes (supplemental Figs. S6, S11, and S12). Conversely, apo-states of the protein deviate further from both holo-states and all deposited crystal structures, and this is particularly manifest in an acidic stretch of the ZA loop (Figs. 6–8). In light of these results, we formulate the following tentative conclusions.

Fast rates are prioritized over ligand specificity

The overall results of the simulations reveal that the ATAD2A-histone H4 complex is highly dynamic and can be

reasonably classified as a fuzzy complex (46) involving two disordered binding partners. This suggests that the bulk of the interaction free energy, which represents a modest affinity, is carried by the acetylated marks while allowing for fast on- and off-rates. This is essential because, for example, the H4Kac5 modification has an estimated half-life of only 2 h in HeLa cells (19). The binding goes beyond the classical typology of conformational selection as both binding partners sample a range of conformations continuously. We hypothesize that a number of these are mutually compatible or are separated from compatible states by low barriers, thus enabling rapid productive encounters.

The primary known function of bromodomains is to recognize and bind acetylated lysine residues. For different acetylation marks, specificity, which is often weak but documented (30, 44, 52, 53), must be conferred by sequence context. Our results let us conclude that the primary design constraint on the interaction surface of ATAD2A is adaptability and not specificity. In both binding modes, the peptide makes numerous and somewhat changeable interactions with ATAD2A, which form a consistent average pattern that is nearly identical for Kac5 and Kac12 (Fig. 3). The adaptability results in a weak coupling between protein conformational states and peptide binding, and the ZA loop of ATAD2A itself can be considered as intrinsically disordered. From ongoing MD simulations, other bromodomains, such as CREBBP, BAZ2A, and BRPF1, preserve this general trend but to a lesser extent. A comprehensive understanding of adaptability to peptide ligands across the bromodomain family is part of ongoing work. This includes the use of chemical probes to expose heterogeneities in binding modes (54–56).

In general, it is not unreasonable that speed is prioritized over affinity to allow fast signaling and adjustments to cellular needs (57). Binding promiscuity is a hallmark of intrinsically disordered proteins and regions (IDPs and IDRs) (58, 59), and the histone code appears to regulate this property for the tails. Promiscuous binding of histone tails to bromodomains is well-documented (8, 30). IDRs often function by transient interactions and are frequently subject to post-translational modifications (60). The properties of IDRs are particularly amenable to the editing and use of these modifications, which first and foremost require accessibility.

Disorder and electrostatics work in concert for recruitment

It is interesting that the ZA loop in its apo-state is able to extend far into solution ($\sim 2 \text{ nm}$ from the binding site for a domain whose long dimension is only $\sim 5 \text{ nm}$), thus providing a clear mechanism for increasing the on-rate. This type of fly-casting mechanism (61, 62) is well-characterized for IDPs, albeit less so for loops in an otherwise folded domain. The increased range is likely to benefit from utilizing long-range attractions, and this is exactly the case for ATAD2A. The most mobile part of the ZA loop is net acidic, and this property may play an important role in “capturing” net positively charged histone tails during binding. This interaction mechanism may be utilized elsewhere, and an example in the literature is the binding of a 17-residue neuropeptide, which is net positively charged, to a highly acidic, 26-residue disordered loop of the

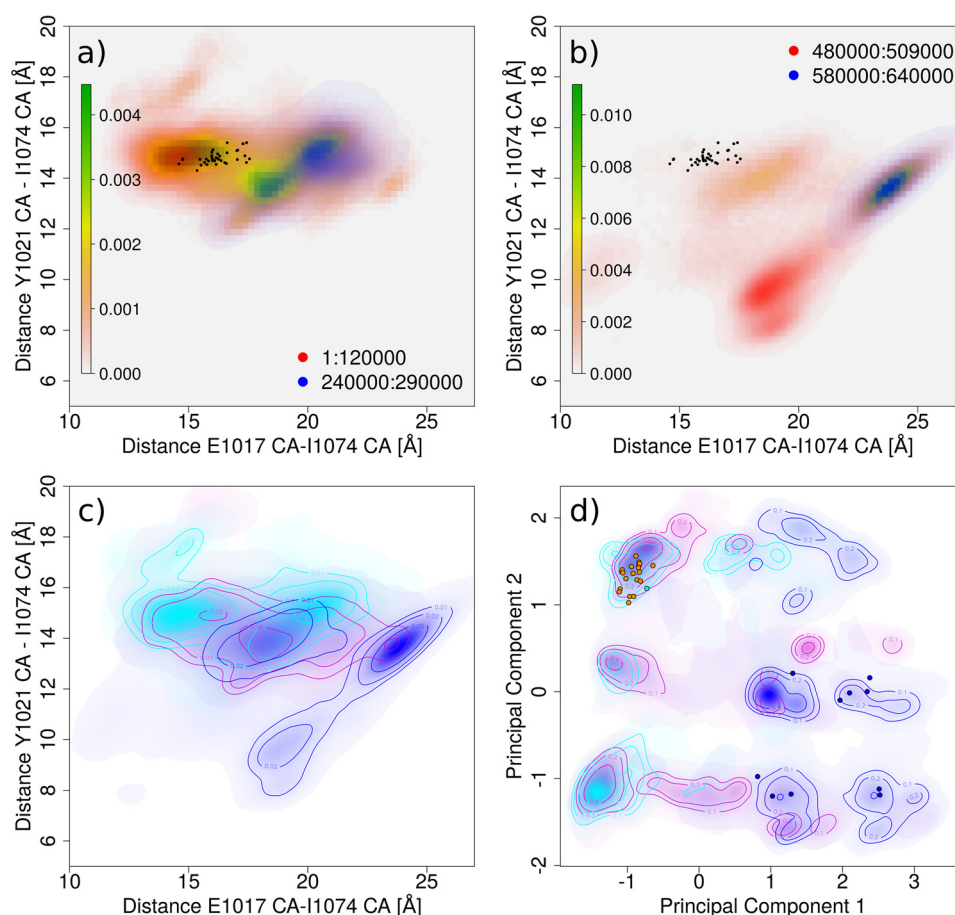


Figure 8. Two-dimensional projections of the data. *a*, histograms in two distances measuring binding pocket aperture for the two sets of holo-simulations. Black dots indicate these values for all crystal structures of ATAD2A released until February 28, 2017. Overlaid color densities distinguish points from main basins in Fig. 6 (see legend to Fig. 6). *b*, this panel is analogous to *a* except the histogram in the background is derived from apo-simulations alone and two different basins were extracted. *c*, an overlay in the same 2D distance space as *a* and *b* of histogram density estimates split into the three simulation groups together with three contour lines per histogram. Apo data are in blue, and holo data with Kac5 and Kac12 inserted are shown in magenta and cyan, respectively. *d*, from sine and cosine values of the ZA loop (residues Lys¹⁰⁰⁴–Asp¹⁰³⁰), we calculated principal components, and the distribution of the first two is shown as a density and contour plot as in *c*. Here, the circles are the starting structures of all of the MD runs and use the same color convention as Fig. 6.

κ -opioid receptor with $\sim 32 \mu\text{M}$ affinity and with little influence on loop conformation as inferred from NMR (63). It should be pointed out that ATAD2A is, as expected, negatively charged under physiological conditions but that the excess charge is located largely at the opposite end of the domain (*i.e.* distal from the binding site in the ~ 25 C-terminal residues) (see Movie S1). This hints at a role of the C-terminal helix in recruitment and/or position of the domain relative to the histone core.

A second intriguing detail emerging from our study is the fact that the histone tail remains disordered upon binding. If the tail were to be bound rigidly to the bromodomain surface, the orientational freedom required for assembling a stable complex between histone, bromodomain, and either DNA or additional proteins (*e.g.* the remainder of the ATAD2 gene product) may be compromised. This is particularly relevant as the histone tails are not that long (*e.g.* 20–25 residues for H4).

Simulations are essential to understand the functional role of disorder

Figs. 6–8 raise the question of the representative value of some crystal structures of ATAD2A, given the exposed nature and high disorder of the ZA loop. Whereas the helix bundle

deviates little from the experimental result in the simulations, the ZA loop in the apo-state explores only “new territory.” It appears as if the holo-simulations recur to the crystal structures, including apo ones, more often, and it can be argued that crystal contacts are the primary determinants of ZA loop conformations in several ATAD2A crystals (see supplemental Fig. S1 and Table 1). The MD simulations have the potential to go beyond the description of the binding mode available from the X-ray structures for three main reasons. First, coordinates are reported only for the first five or six histone residues (*i.e.* 9–13, 3–7, and 1–6) for the complexes with H4(9–16), H4(3–16), and H4(1–20) (PDB codes 4QUT, 4QUU, and 4TT2, respectively). This is expected and consistent with the flexibility observed in the simulations (*e.g.* schematics in Fig. 2). Second, truncation of constructs can also lead to spurious electrostatic interactions because of charged N or C termini. As an example, in the structure of ATAD2A with the H4(9–16) peptide segment (PDB code 4QUT), the salt bridges between the positively charged amino group of G9 and the aspartate triad on the BC loop are probably artifactual. Third, as shown in prior work, the Kac mark can populate alternative binding modes, which differ in insertion depth, hydrogen-bonded partners, the number of

water molecules in the binding site, etc. We find this to be the case also for ATAD2A as [supplemental Fig. S7](#) documents. This includes bound states where the Kac binding pocket itself is severely, although reversibly, deformed. Thus, whereas the large number of crystal structures lets Kac binding to bromodomains appear as a homogeneous process with a well-conserved, ordered reference state, the simulations highlight the critical role that disorder plays in this binding, and we argue here that it confers properties that are necessary for function.

Fuzzy interactions may promote cooperativity between nearby marks

In the binding of histone tails to bromodomains, the immediate sequence context can be similar or identical. For example, in human histone H4, the sequence GKacG is encountered for both positions 5 and 12 (and Kac8 would be the same). Given that these motifs can occur multiple times along the sequence, we conjecture here not only that they can bring two different binding partners close to each other in a ternary complex, but that they are also read and understood by a single domain although only a single mark is bound at any given time. A quantitative understanding of the thermodynamic and kinetic implications of multiple acetylations, which are well-known experimentally (e.g. for ATAD2A and histone H4 (44) or BAZ2A and histone H2A (30)) will require a different simulation strategy, most likely relying on enhanced sampling techniques. These are required because the unbinding time exceeds 1 μ s, which makes simulations of reversible dissociation and association prohibitively expensive by conventional sampling. The high prevalence of intrinsic disorder in proteins involved in transcriptional regulation (64) makes it seem probable that the recognition of modified histone tails in fuzzy complexes is a general mechanism for this cellular process.

Conclusions

The regulation of transcription (and many bromodomains are thought or known to be involved in this) is a surprisingly stochastic process due to the low copy numbers involved (see Introduction). Nonetheless, proper cell function necessitates outcomes that are highly predictable despite this stochasticity (29). We have reported here the binding of histone tail peptides in fuzzy complexes, which allude to three distinct roles that disorder can play in achieving these predictable outcomes. First, disorder can enable binding promiscuity, which leads to functional redundancy in the regulatory network, and this provides robustness. Second, the similarity of average contact patterns suggests a dynamic switching that would enable cooperativity between nearby marks. Such cooperativity would allow an ultrasensitive fine-tuning of the recruitment of bromodomains to histones. Third, the absence of any disorder-to-order transition during binding is likely to impart speed to complex formation, and fast binding and unbinding rates avoid unnecessary lag during cellular responses to regulatory signals. It has not been easy to understand the outcomes of interfering with bromodomain function (31), and we believe that the results of this study provide clues regarding the origins of this difficulty.

Experimental procedures

The protocols for preparing the system, running the MD simulations, and analyzing the data are described briefly here. Additional methodological details are provided as [supplemental material](#) ([supplemental Methods](#) and [Table S1](#)).

Simulation protocols

The coordinates of the ATAD2A bromodomain in complex with the histone peptides H4(3–16)K5ac and H4(9–16)K12ac were downloaded from the Protein Data Bank (PDB codes 4QUU and 4QUT (37), respectively). The sequence of the N-terminal residues 1–16 of the H4 peptide is SGRGKGGKGLGKGGAK (H4(1–16)). All MD simulations were carried out with the diacetylated H4(1–16)K5acK12ac peptide, and the C-terminal residue (*i.e.* Lys¹⁶) was capped with *N*-methylamide to emulate the linkage to the rest of the histone H4.

Starting from the X-ray structure, the peptide residues that were not found in the crystal were manually built using PyMOL (PyMOL Molecular Graphics System, version 1.7.4, Schrödinger, LLC, New York) and relaxed through a broad combination of Monte Carlo moves as implemented in the CAMPARI software package (<http://campari.sourceforge.net>),⁴ keeping the bromodomain frozen and restraining the resolved peptide residues in the binding pocket (see [supplemental Methods](#) and [Table S1](#) in the [supplemental material](#)). 10 evenly spaced snapshots were extracted from the run of 10⁶ elementary moves. All of the crystal waters not clashing with the added residues of the peptide were kept. A short NPT and NVT water equilibration was run before the actual production at 310 K in the NVT ensemble.

For either Kac5 or Kac12 inserted, 10 peptide-bound (holo) simulations were run for 0.55 μ s each. Simulations of the free (apo) ATAD2A bromodomain were started from the structure with PDB code 3DAI (8). To generate somewhat diverse starting conformations, we ran an initial simulation from the crystal structure (after equilibration), from which, after the first 120 ns, we selected one structure every 20 ns to be used as a starting point for 10 additional independent runs of 0.37 μ s each. The first simulation was run for 0.43 μ s and is included in the analysis. Last, simulations of the free peptide used a set of 10 starting conformations diverse in size and shape and were run for 100 ns each. Detailed settings are provided as [supplemental material](#) ([supplemental Methods](#)), and a pictorial summary of all of the MD runs carried out is shown in [Fig. 1c](#).

Data analysis

Simple analyses are described in the figure legends, and additional details can be found in the [supplemental material](#). To generate a comprehensive view of the thermodynamics and kinetics of the simulated system, we employed the so-called progress index algorithm (49). In it, snapshots are arranged in such a way that at each step the (geometrically) closest snapshot to any of the prior snapshots in the sequence is added. We rely on a scalable approximate implementation as published. This requires a preorganization step by means of a tree-based clus-

⁴ Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.

tering (65). The starting point was always picked as the centroid representative of the largest cluster in this auxiliary clustering. The idea behind the progress index is that geometrically homogeneous states are all resolved without overlap and that the resolution is maximal (snapshot level).

Different annotations of the trajectory can be plotted per snapshot as arranged by the progress index to obtain a SAP-PHIRE plot (48), which helps in the interpretation of the discovered basins. The cut function we show here for kinetic partitioning is the so-called local cut in the original publication (49). It corresponds to the inverse of the mean first passage time between two states defined as the 70,000 snapshots to the left or right of the current position in the progress index. As plotted, this annotation creates a pseudo-free energy profile (*i.e.* low values within basins and high values at barriers). The particular choice for the width of the two states is expected to be sufficient as long as no basins are significantly larger than twice this value. As the distance function for generating the progress index, we used either the RMSD of selected atoms along the ZA and BC loops after alignment of the C α atoms of α -helices or the dihedral angle distance (65) of the ϕ and ψ angles of the ZA loop (see supplemental material).

Author contributions—A. V. and A. C. designed the study, and all authors contributed to the writing of the manuscript. C. L. performed the simulations and analyzed the data. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgment—We thank Jean-Rémy Marchand for interesting discussions.

References

- Zeng, L., and Zhou, M.-M. (2002) Bromodomain: an acetyl-lysine binding domain. *FEBS Lett.* **513**, 124–128
- Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M.-M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496
- Owen, D. J., Ornaghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A. A. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *EMBO J.* **19**, 6141–6149
- Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. *Nature* **403**, 41–45
- Turner, B. M. (2000) Histone acetylation and an epigenetic code. *Bioessays* **22**, 836–845
- Rando, O. J. (2012) Combinatorial complexity in chromatin structure and function: revisiting the histone code. *Curr. Opin. Genet. Dev.* **22**, 148–155
- Filippakopoulos, P., and Knapp, S. (2012) The bromodomain interaction module. *FEBS Lett.* **586**, 2692–2704
- Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., Gingras, A.-C., Arrowsmith, C. H., and Knapp, S. (2012) Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**, 214–231
- Marchand, J.-R., and Caflisch, A. (2015) Binding mode of acetylated histones to bromodomains: variations on a common motif. *Chem. Med. Chem.* **10**, 1327–1333
- Steiner, S., Magno, A., Huang, D., and Caflisch, A. (2013) Does bromodomain flexibility influence histone recognition? *FEBS Lett.* **587**, 2158–2163
- Jain, A. K., and Barton, M. C. (2017) Bromodomain histone readers and cancer. *J. Mol. Biol.* **429**, 2003–2010
- Xu, Y., Zhang, S., Lin, S., Guo, Y., Deng, W., Zhang, Y., and Xue, Y. (2017) WERAM: a database of writers, erasers and readers of histone acetylation and methylation in eukaryotes. *Nucleic Acids Res.* **45**, D264–D270
- Sanchez, R., Meslamani, J., and Zhou, M.-M. (2014) The bromodomain: from epigenome reader to druggable target. *Biochim. Biophys. Acta* **1839**, 676–685
- Dey, A., Chitsaz, F., Abbasi, A., Misteli, T., and Ozato, K. (2003) The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 8758–8763
- Phair, R. D., Scaffidi, P., Elbi, C., Vecerová, J., Dey, A., Ozato, K., Brown, D. T., Hager, G., Bustin, M., and Misteli, T. (2004) Global nature of dynamic protein-chromatin interactions *in vivo*: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. *Mol. Cell. Biol.* **24**, 6393–6402
- Mazza, D., Abernathy, A., Golob, N., Morisaki, T., and McNally, J. G. (2012) A benchmark for chromatin binding measurements in live cells. *Nucleic Acids Res.* **40**, e119–e119
- Hager, G. L., McNally, J. G., and Misteli, T. (2009) Transcription dynamics. *Mol. Cell* **35**, 741–753
- Steffen, P. A., Fonseca, J. P., and Ringrose, L. (2012) Epigenetics meets mathematics: towards a quantitative understanding of chromatin biology. *Bioessays* **34**, 901–913
- Zheng, Y., Thomas, P. M., and Kelleher, N. L. (2013) Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites. *Nat. Commun.* **4**, 2203
- Barth, T. K., and Imhof, A. (2010) Fast signals and slow marks: the dynamics of histone modifications. *Trends Biochem. Sci.* **35**, 618–626
- Schreiber, G., Haran, G., and Zhou, H. X. (2009) Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* **109**, 839–860
- Schreiber, G., and Fersht, A. R. (1996) Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.* **3**, 427–431
- Kiel, C., Selzer, T., Shaul, Y., Schreiber, G., and Herrmann, C. (2004) Electrostatically optimized Ras-binding Ral guanine dissociation stimulator mutants increase the rate of association by stabilizing the encounter complex. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9223–9228
- Blöchliger, N., Xu, M., and Caflisch, A. (2015) Peptide binding to a PDZ domain by electrostatic steering via nonnative salt bridges. *Biophys. J.* **108**, 2362–2370
- Peng, Z., Mizianty, M. J., Xue, B., Kurgan, L., and Uversky, V. N. (2012) More than just tails: intrinsic disorder in histone proteins. *Mol. Biosyst.* **8**, 1886–1901
- Hansen, J. C., Lu, X., Ross, E. D., and Woody, R. W. (2006) Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J. Biol. Chem.* **281**, 1853–1856
- Mittag, T., Kay, L. E., and Forman-Kay, J. D. (2010) Protein dynamics and conformational disorder in molecular recognition. *J. Mol. Recognit.* **23**, 105–116
- Fuxreiter, M. (2012) Fuzziness: linking regulation to protein dynamics. *Mol. Biosyst.* **8**, 168–177
- McAdams, H. H., and Arkin, A. (1999) It's a noisy business! Genetic regulation at the nanomolar scale. *Trends Genet.* **15**, 65–69
- Philpott, M., Yang, J., Tumber, T., Fedorov, O., Uttarkar, S., Filippakopoulos, P., Picaud, S., Keates, T., Felletar, I., Ciulli, A., Knapp, S., and Heightman, T. D. (2011) Bromodomain-peptide displacement assays for interaction mapping and inhibitor discovery. *Mol. Biosyst.* **7**, 2899–2908
- Shi, J., and Vakoc, C. R. (2014) The mechanisms behind the therapeutic activity of BET bromodomain inhibition. *Mol. Cell* **54**, 728–736
- Duan, Z., Zou, J. X., Yang, P., Wang, Y., Borowsky, A. D., Gao, A. C., and Chen, H.-W. (2013) Developmental and androgenic regulation of chromatin regulators EZH2 and ANCCA/ATAD2 in the prostate via MLL histone methylase complex. *Prostate* **73**, 455–466
- Revenko, A. S., Kalashnikova, E. V., Gemo, A. T., Zou, J. X., and Chen, H.-W. (2010) Chromatin loading of E2F-MLL complex by cancer-associated coregulator ANCCA via reading a specific histone mark. *Mol. Cell. Biol.* **30**, 5260–5272
- Caron, C., Lestrat, C., Marsal, S., Escoffier, E., Curtet, S., Virolle, V., Barbry, P., Debernardi, A., Brambilla, C., Brambilla, E., Rousseaux, S., and Khochbin, S. (2010) Functional characterization of ATAD2 as a new can-

- cer/testis factor and a predictor of poor prognosis in breast and lung cancers. *Oncogene* **29**, 5171–5181
35. Ciró, M., Prosperini, E., Quarto, M., Grazini, U., Walfridsson, J., McBlane, F., Nucifero, P., Pacchiana, G., Capra, M., Christensen, J., and Helin, K. (2009) ATAD2 is a novel cofactor for MYC, overexpressed and amplified in aggressive tumors. *Cancer Res.* **69**, 8491–8498
36. Poncet-Montange, G., Zhan, Y., Bardenhagen, J. P., Petrocchi, A., Leo, E., Shi, X., Lee, G. R., 4th, Leonard, P. G., Geck Do, M. K., Cardozo, M. G., Andersen, J. N., Palmer, W. S., Jones, P., and Ladbury, J. E. (2015) Observed bromodomain flexibility reveals histone peptide- and small molecule ligand-compatible forms of ATAD2. *Biochem. J* **466**, 337–346
37. Morozumi, Y., Boussouar, F., Tan, M., Chaikuad, A., Jamshidikia, M., Colak, G., He, H., Nie, L., Petosa, C., de Dieuleveult, M., Curtet, S., Vitte, A.-L., Rabatel, C., Debernardi, A., Cosset, F.-L., *et al.* (2016) Atad2 is a generalist facilitator of chromatin dynamics in embryonic stem cells. *J. Mol. Cell. Biol.* **8**, 349–362
38. Harner, M. J., Chauder, B. A., Phan, J., and Fesik, S. W. (2014) Fragment-based screening of the bromodomain of ATAD2. *J. Med. Chem.* **57**, 9687–9692
39. Chaikuad, A., Petros, A. M., Fedorov, O., Xu, J., and Knapp, S. (2014) Structure-based approaches towards identification of fragments for the low-druggability ATAD2 bromodomain. *Med. Chem. Commun.* **5**, 1843–1848
40. Bamborough, P., Chung, C.-W., Furze, R. C., Grandi, P., Michon, A.-M., Sheppard, R. J., Barnett, H., Diallo, H., Dixon, D. P., Douault, C., Jones, E. J., Karamshi, B., Mitchell, D. J., Prinjha, R. K., Rau, C., *et al.* (2015) Structure-based optimization of naphthyridones into potent ATAD2 bromodomain inhibitors. *J. Med. Chem.* **58**, 6151–6178
41. Demont, E. H., Chung, C.-W., Furze, R. C., Grandi, P., Michon, A.-M., Wellaway, C., Barrett, N., Bridges, A. M., Craggs, P. D., Diallo, H., Dixon, D. P., Douault, C., Emmons, A. J., Jones, E. J., Karamshi, B. V., *et al.* (2015) Fragment-based discovery of low-micromolar ATAD2 bromodomain inhibitors. *J. Med. Chem.* **58**, 5649–5673
42. Magno, A., Steiner, S., and Cafilisch, A. (2013) Mechanism and kinetics of acetyl-lysine binding to bromodomains. *J. Chem. Theory Comput.* **9**, 4225–4232
43. Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T., and Allis, C. D. (1995) Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1237–1241
44. Koo, S. J., Fernández-Montalván, A. E., Badock, V., Ott, C. J., Holton, S. J., von Ahsen, O., Toedling, J., Vittori, S., Bradner, J. E., and Górnáncz, M. (2016) ATAD2 is an epigenetic reader of newly synthesized histone marks during DNA replication. *Oncotarget* **7**, 70323–70335
45. Potoyan, D. A., and Papoian, G. A. (2012) Regulation of the H4 tail binding and folding landscapes via Lys-16 acetylation. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17857–17862
46. Tompa, P., and Fuxreiter, M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **33**, 2–8
47. Convertino, M., Vitalis, A., and Cafilisch, A. (2011) Disordered binding of small molecules to A β (12–28). *J. Biol. Chem.* **286**, 41578–41588
48. Blöchliger, N., Vitalis, A., and Cafilisch, A. (2014) High-resolution visualization of the states and pathways sampled in molecular dynamics simulations. *Sci. Rep.* **4**, 6264
49. Blöchliger, N., Vitalis, A., and Cafilisch, A. (2013) A scalable algorithm to order and annotate continuous observations reveals the metastable states visited by dynamical systems. *Comput. Phys. Commun.* **184**, 2446–2453
50. Kabsch, W., and Sander, C. (1983) Dictionary of protein Secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637
51. Receveur-Bréchet, V., Bourhis, J.-M., Uversky, V. N., Canard, B., and Longhi, S. (2006) Assessing protein disorder and induced folding. *Proteins* **62**, 24–45
52. Kanno, T., Kanno, Y., Siegel, R. M., Jang, M. K., Lenardo, M. J., and Ozato, K. (2004) Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol. Cell* **13**, 33–43
53. Hassan, A. H., Awad, S., Al-Natour, Z., Othman, S., Mustafa, F., and Rizvi, T. A. (2007) Selective recognition of acetylated histones by bromodomains in transcriptional co-activators. *Biochem. J.* **402**, 125–133
54. Marchand, J.-R., Lolli, G., and Cafilisch, A. (2016) Derivatives of 3-amino-2-methylpyridine as BAZ2B bromodomain ligands: *in silico* discovery and *in crystallo* validation. *J. Med. Chem.* **59**, 9919–9927
55. Unzue, A., Zhao, H., Lolli, G., Dong, J., Zhu, J., Zechner, M., Dolbois, A., Cafilisch, A., and Nevado, C. (2016) The “gatekeeper” residue influences the binding mode of acetyl indoles to bromodomains. *J. Med. Chem.* **59**, 3087–3097
56. Zhu, J., and Cafilisch, A. (2016) Twenty crystal structures of bromodomain and PHD finger containing protein 1 (BRPF1)/ligand complexes reveal conserved binding motifs and rare interactions. *J. Med. Chem.* **59**, 5555–5561
57. Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., and Obradović, Z. (2002) Intrinsic disorder and protein function. *Biochemistry* **41**, 6573–6582
58. Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M., and Uversky, V. N. (2005) Flexible nets. *FEBS J.* **272**, 5129–5148
59. Tompa, P., Schad, E., Tantos, A., and Kalmar, L. (2015) Intrinsically disordered proteins: emerging interaction specialists. *Curr. Opin. Struct. Biol.* **35**, 49–59
60. Wright, P. E., and Dyson, H. J. (2015) Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29
61. Shoemaker, B. A., Portman, J. J., and Wolynes, P. G. (2000) Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 8868–8873
62. Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021–1025
63. Björnerås, J., Kurnik, M., Oliveberg, M., Gräslund, A., Mäler, L., and Danielsson, J. (2014) Direct detection of neuropeptide dynorphin A binding to the second extracellular loop of the κ opioid receptor using a soluble protein scaffold. *FEBS J.* **281**, 814–824
64. Minezaki, Y., Homma, K., Kinjo, A. R., and Nishikawa, K. (2006) Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation. *J. Mol. Biol.* **359**, 1137–1149
65. Vitalis, A., and Cafilisch, A. (2012) Efficient construction of mesostate networks from molecular dynamics trajectories. *J. Chem. Theory Comput.* **8**, 1108–1120

The ATAD2 bromodomain binds different acetylation marks on the histone H4 in similar fuzzy complexes

Cassiano Langini, Amedeo Caflisch and Andreas Vitalis

J. Biol. Chem. 2017, 292:16734-16745.

doi: 10.1074/jbc.M117.786350 originally published online August 10, 2017

Access the most updated version of this article at doi: [10.1074/jbc.M117.786350](https://doi.org/10.1074/jbc.M117.786350)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

Supplemental material:

<http://www.jbc.org/content/suppl/2017/08/10/M117.786350.DC1>

This article cites 65 references, 13 of which can be accessed free at

<http://www.jbc.org/content/292/40/16734.full.html#ref-list-1>